Available online at www.sciencedirect.com

ScienceDirect

Speech Communication 57 (2014) 13–38

SPEECH COMMUNICATION

www.elsevier.com/locate/specom

An educational platform to demonstrate speech processing techniques on Android based smart phones and tablets

Roger Chappel *, Kuldip Paliwal

Signal Processing Laboratory, School of Engineering, Griffith University, Nathan Campus, Brisbane QLD 4111, Australia

Received 7 November 2012; received in revised form 29 July 2013; accepted 6 August 2013
Available online 19 August 2013

Abstract

This work highlights the need to adapt teaching methods in digital signal processing (DSP) on speech to suit shifts in generational learning behavior, furthermore it suggests the use of integrating theory into a practical smart phone or tablet application as a means to bridge the gap between traditional teaching styles and current learning styles. The application presented here is called "Speech Enhancement for Android (SEA)" and aims at assisting in the development of an intuitive understanding of course content by allowing students to interact with theoretical concepts through their personal device. SEA not only allows the student to interact with speech processing methods, but also enables the student to interact with their surrounding environment by recording and processing their own voice. A case study on students studying DSP for speech processing found that by using SEA as an additional learning tool enhanced their understanding and helped to motivate students to engage in course work by way of having ready access to interactive content on a hand held device. This paper describes the platform in detail acting as a road-map for education institutions, and how it can be integrated into a DSP based speech processing education framework.

Keywords: Digital signal processing (DSP); Electronic engineering education; Speech enhancement; Android; Smart phone

1. Introduction

Teaching digital signal processing (DSP) is often a requirement in the undergraduate and post graduate electronic engineering curriculum.

Students undertaking this field of study cover topics such as: Discrete Fourier transform (DFT) (Nussbaumer, 1981) and spectral estimation and filter design (Talor, 1983). At Griffith university, DSP is currently taught in three stages. Firstly, the continuous time analysis of signals and systems (Soliman and Srinath, 1990). Secondly, the introduction of discrete time signals and systems incorporating digital filtering on stationary processes (Proakis and Manolakis, 2007). And, lastly advanced topics such as spectral estimation and adaptive filtering for stochastic signals (Haykin, 1991; Kay, 1993; Hayes, 1996).

Each stage is accompanied by a series of laboratory sessions, exposing the student to programming techniques which can be useful in the industry, both on a simulation environment such as Matlab (Mitra, 2006) and on a hardware based tool such as the Texas Instruments 6713 digital signal processor starter kit (DSK) (Chassaing, 2002). The practical element is often concluded with a primary project providing a valuable opportunity for students to apply prior DSP learnings to a known real world application in areas such as speech coding for telecommunication (Crochiere, 1981) and automatic speaker verification (Rosenberg, 1976). In this paper, we propose an additional tool which we can be incorporated into the DSP syllabus to aid in demonstrating fundamental principals both during laboratories and lectures along side Matlab, DSK and relevant text books. This tool is called "Speech Enhancement

* Corresponding author. Tel.: +61 434574285.
E-mail address: roger.chappel@girifithuni.edu.au (R. Chappel).
URL: http://www.enhancementapp.com (R. Chappel).

0167-6393/$ - see front matter
http://dx.doi.org/10.1016/j.specom.2013.08.002

for Android" (SEA) and has been developed on the Google Android platform with the intention to bring DSP principals to the increasingly popular smart phones, which now account for 52.5% of the global smart phone market share, double that of 2010 (Gartner, 2011b). Additionally, SEA has been developed to support a wide range of Android based tablets as they begin to gain momentum in the global market (Gartner, 2011a; Milanesi, 2011). The mobility and accessibility of personal hand held devices over laptops in *K*-12 education has been assessed, resulting in an increase in student engagement (Soloway et al., 2001) and an increase in effective technical education in mathematics (Yerushalmy and Ben-Zaken, 2004).

Recent advances in low-powered CPU design have introduced a new era of high performance hand held devices, capable of intensive graphics rendering with on board graphics processing units (GPU) and complex operations using multi-threaded and multi-core technology. NVIDIA (2010), ARM (2009). As the popularity is increasing and manufacturing methods are improving, these devices are now becoming accessible to students. This is opening many opportunities for education institutions to bring low cost interactive educational tools to every student. SEA capitalises on these rising opportunities by utilizing the computing power of modern hand-held devices which otherwise would have only been available on a computer.

The objective of this tool is to provide additional flexibility in the delivery of course content to accommodate for shifts in generational learning styles and behaviors (Caillaud and Cohen, 2000; Billings et al., 2005; Rosenbaum and Rochford, 2008; Salajan et al., 2010). In addition to shifts occurring in learning behaviors, other factors such as technology trends being closely associated with lifestyle is becoming prevalent (Wellman, 2002; Hassanlou et al., 2009; Martin et al., 2011). The majority of students in established universities will have access to a smart phone which provides a portal, integrating individuals' reality with that of their online virtual-reality; where study, work, entertainment and social life amalgamate into one unifying device (Donath and Boyd, 2004; Ellison et al., 2007; Humphreys, 2010; Song et al., 2012). Studies show (White, 1989; Price et al., 2003; O'NeilJr and Perez, 2003; Lubis et al., 2010; Price and Falcao, 2011), that adapting content to suit these trends can encourage and motivate students to engage in activities and enhance their learning potential. Moreover, integrating content into preferred communication technologies which appeal to each generation through visual and interactive means may also influence personal intrinsic learning and cognitive skills which could accelerate student's understanding and comprehension of content (Sims, 1997; Joiner et al., 2006; Kong et al., 2012; Rienties et al., 2009).

SEA focuses on speech processing methods, not only introducing the student to a real world application for DSP such as the field of speech enhancement but also demonstrating the principals of analysing a human speech signal. These principals address issues such as stationarity, spectral estimation, the role of the magnitude and phase spectrum, filtering, statistical methods for noise removal and the spectrogram representation of speech. These topics extend beyond that of DSP and can be useful in other disciplines such as linguistics, speech pathology and audiology.

Another goal of this tool is to present many sophisticated signal processing techniques through a simple user interface (UI) allowing a realizable and practical interpretation of every element. It is designed to give the user complete control of all configurable settings, allowing the abstraction of the more advanced speech enhancement techniques to a simple, intuitive example of each concept. This allows researchers in the field of speech enhancement to examine configuration changes rapidly, saving them time on programming and interpretation. Results can be saved and shared via e-mail or any other supported file transfer method.

This paper is organised as follows. A review of current mobile DSP applications and a description of the development process of SEA are shown in Section 1 and 2. The primary elements and functionality of SEA are discussed in Section 4. Section 5 describes the short-time analysis–modification–synthesis (AMS) framework which is used for the speech processing in this work. Section 6 illustrates the procedure undertaken to record and display the spectrogram of a speech signal. Section 7 demonstrates the short-time analysis function of SEA, allowing the user to view the time and frequency domain representation of each frame. Section 7 describes the linear prediction, source-filter separation method used in SEA. Section 10 illustrates the procedure undertaken to convey the role of the short-time magnitude and phase spectra on speech intelligibility. Section 11 describes the speech enhancement principals behind the implementation within SEA. Section 12 details a subjective test evaluating SEA's effectiveness as a learning tool in a learning environment. Section 16 explains future objectives and challenges faced for the SEA project. And lastly, Section 17 contains a summary with concluding comments.

2. Related mobile platforms

The concept of mobile educational platforms in electrical engineering and bioinfomatics has been generating a lot of interest since the evolution of the smartphone (Teng and Helps, 2010; Finkelstein et al., 2010). Prior to the smartphone revolution, web-based learning aids were the central theme on providing accessibility to course content for students at home or on their own laptop while on campus through an extensive wireless network. These applications ranged from complete online courses and laboratories (Maiti et al., 2011) to simple Java applets to accompany other course content (Gonvalves and Canesin, 2001). Several online DSP learning aids have been developed (Ko et al., 2003; Jackson et al., 2001), however, only a small minority have extended their platforms to